DOCUMENT RESUME

ED 435 682 TM 030 313

AUTHOR Kirchner, Robert A.
TITLE Creating and Sustaining State and Local Evaluation Capacities.
PUB DATE 1999-11-00
NOTE 20p.; Paper presented at the Annual Meeting of the American Evaluation Association (Orlando, FL, November 3-6, 1999).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Methods; Local Government; Policy Formation; *Program Evaluation; *State Programs; *Systems Approach

ABSTRACT

Assessment and evaluation can be important tools in improving the effectiveness and quality of programs, especially criminal justice programs, if they are integrated into the plans and designs of strategies instead of being added later. An evaluation systems approach to program development is needed to ensure that effectiveness and efficiency are maintained and that progress on program objectives is communicated to policy makers, managers, and the public. There are four criteria that define an effectively managed program: (1) acceptable description of goals and objectives; (2) linkage between program activities and objectives; (3) performance information; and (4) acceptable performance. An analytical model is suggested that demonstrates the linear progress of evaluation and performance data from program goals to specific activities, through performance measurement, to analysis and interpretation of the results and necessary corrective actions. Evaluation is an important program development tool, not an isolated event that can be performed once and then forgotten. Three appendixes contain charts about evaluation systems and models. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 435 682

CREATING AND SUSTAINING STATE AND LOCAL EVALUATION CAPACITIES[1]

Robert A. Kirchner, Ph.D.
Senior Advisor for Evaluation
Bureau of Justice Assistance
U.S. Department of Justice

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
R. Kirchner
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Introduction.

A need exists for information on programs that work and confirmation on why they work. Criminal justice programs, and specifically new cost-effective, innovative demonstration activities must include research and evaluation components before implementation begins. We all work in complex political and bureaucratic environments attempting to accomplish numerous and conflicting goals, and having critical data on program performance is often the best offense for program managers.

Assessment and evaluation can be important tools in improving the effectiveness and quality of the success of criminal justice programs, if they are integrated into the plans and designs of criminal justice strategies rather than added on afterward. An evaluation systems approach to program development is needed to ensure that: (1) effectiveness and efficiency are maintained within the program and (2) progress on the program objectives is communicated to key policy makers, managers, and the public. Building evaluation into program development truly makes a program "whole" and ready for implementation. Evaluation activities and methods become "applied" in their use, and provide for on-going program improvement and - perhaps more important - opportunities for future, more intensive evaluation. "Applied" simply dictates that evaluations answer the questions of decision makers, and present results and recommendations that are readily useful to practitioners.

Program Management, Program Success and Program Justification

The application of the framework and methods below creates an analytical model that must be made an ongoing part of the program management for drug courts. Someone within program management, for instance, must be assigned responsibility for conducting periodic effectiveness evaluations pursuant to the model. Additionally, the model is designed only to highlight accomplishments or shortcomings in program performance, not their causes or solutions. Positive findings should be followed up on to determine if program replication is

TM030313

---

[1]Prepared for presentation at the Annual Conference of the American Evaluation Association, *Evaluation '99 - The Territory Ahead: Foundations and Frontiers*, held in Orlando FL, November 4-7, 1999. The session was, "Program Evaluation Tools for the New Century."

appropriate. Program deficiencies will have to be followed up on to isolate causes and ensure that corrective actions are taken. With information provided by program evaluations, program managers can ensure effective program performance and fix program accountability in the eyes of Federal, State or local officials, legislative oversight, and the public.

Evaluation of program performance should be accomplished on a continuing basis and should provide an overall framework for all program participants to benefit by utilizing evaluation findings and recommendations. An attempt should be made to apply the concepts, principles, and techniques embodied in evaluation literature to fit the characteristics and peculiarities of specific public programs. Common sense in designing evaluations, however, is important when you cannot do random assignment or compare control groups. In the end, methods should meet the needs of programs, because programs cannot always meet the criteria of rigorous evaluation research.

If program evaluation is to be a fundamental part of effective public administration, the primary responsibility for program evaluation should rest with oversight officials. To more closely integrate program evaluation and program administration, we propose that program managers be assigned primary responsibility for assuring that program evaluation functions: (1) demonstrate to the satisfaction of oversight officials, the extent to which the program is effectively administered, and (2) support the program manager in producing an effectively administered program.

Public sector system activities are diverse. For criteria to apply across programs, they should not relate to more than specific types or classes of programs, organizational arrangements, and management styles. Consequently, the criteria are described in terms of the program management and accountability processes. These two processes define, on a case-by-case basis for a program, what the program is and how its effectiveness is to be judged.

Evaluation is a powerful tool for planning, developing and managing criminal justice programs. As an objective means of documenting success, identifying programs and guiding refinements, program evaluation is important to a variety of stakeholders.

A useful definition for evaluation: Evaluation involves the systematic assessment of whether and to what extent projects or programs are implemented as intended and whether they achieve their intended objectives. This entails asking questions about programs, and collecting and analyzing information to learn about program operations and to discover program results.

Program managers need this information to guide program development and to demonstrate success. Policy makers and funding sources need it to identify what works and where to focus resources. The expansion, contraction, elimination and modification of programs are often influenced by evaluation findings.

There are four criteria which define an effectively managed program[2]. They define a set of ideal conditions for program implementation and performance -- conditions to be brought about through the management and accountability processes. Since the criteria describe the ideal, the relevant question is, "To what extent do programs achieve the four criteria?"

A brief description of the FOUR criteria follows:

1. **Acceptable Description of Goals and Objectives**: Goals, the end results that programs pursue, are realistic and clearly stated. Program objectives (the effects or results to be achieved by the program in pursuing its goals) are both measurable and achievable.

2. **Linkage between Program Activities and Objectives**: The program has sufficient and appropriate activities in place to achieve the objective (results) expected by program managers. "Sufficient and appropriate activities" means there is evidence that the existing pattern of program activities can produce the results expected. In other words, the causal linkage between program activities and objectives is plausible.

3. **Performance Information**: Performance measures are developed which signal whether and/or to what extent the program is meeting its objectives (achieving expected results). This information is obtained by measuring the program's actual results, then comparing them with the program's expected results.

4. **Acceptable Performance**: The program meets or exceeds the expectations (objectives) set for it, and its actual performance is acceptable to program managers and oversight officials. This criterion recognizes there may be times when a program does not fully achieve its objectives (due to unforeseen and uncontrollable events), but is nevertheless considered to be performing successfully.

Rushing to Implementation

The need to focus on innovative programs which are successfully contributing to the Nation's efforts to develop effective and efficient programs, while solving societal problems, is well recognized. The study of individual programs, located across the country, will help identify what works, and what may be transferrable to other locations. Unfortunately, too often the programs identified were put in place to react to immediate demands and their "rush to implementation" did not include building in agreements for assessment and evaluation.

---

[2]Appendix A, "How Evaluation Fits In," complements the discussion of the four criteria.

Incremental phases of developing a complete evaluation system.[3] Assessing the effectiveness of programs for the purpose of finding out how well they have been implemented and to compare the extent to which the activities funded have achieved the program's goals is always possible, even given limited resources and funding constraints. The products of an evaluation system are designed to provide administrators and policy makers with an improved understanding of whether specific activities accomplish their desired results of enhancing the effectiveness of the stated program strategies.

"Program evaluation" is defined as follows: A systematic assessment of the results or outcomes of program efforts to measure actual outcomes against the intended outcomes of the program; to discover achievement and results; to discover deviations from planned achievements; to judge the worth of the program; to identify unintended consequences; and to recommend expansion, contraction, elimination, or modification of the program.

It is obvious from the preceding definition that program evaluation is an invaluable aid in planning, developing, and managing programs. To be effective, however, program evaluation efforts must be placed within the broader context of program management. A flexible capacity for internal self-evaluation is fundamental to the management and ongoing improvement of programs.

The three incremental phases envisioned for a **"complete"** evaluation system are: program logic analysis; process evaluation and impact/intensive evaluation[4].

The purpose of **"program logic analysis"** is to provide the basic foundation of program design, including the established linkages between objectives and program activities and consensus on performance and impact indicators. Developing a "model" of the program in the planning stage permits managers to formulate their expectations for program outcomes, which can used later for program analysis and evaluation. Actual results are of little use, if they cannot be compared with expected results.

---

[3] The approach and definitions presented here are fully explained and demonstrated in: Kirchner, Robert A., Roger K. Przybylski and Ruth A. Cardella Assessing the Effectiveness of Criminal Justice Programs. Assessment and Evaluation Handbook Series Number 1, January 1994. U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Assistance. This publication is available on the INTERNET at: **www.bja.evaluationwebsite.org**. Also see Appendix B, "Complete Evaluation System," for an overview, approaches and incremental phases of the evaluation process.

[4] See Appendix C, "Program Evaluation: A Realistic Framework," which includes specific examples.

Program logic analysis also provides needed information to conduct initial implementation studies and assessments directed by inquiries on program progress. Implementation analysis specifically identifies what is often forgotten -- descriptions of program activities, which are the formative basis for evaluation. Hence, the program manager can begin with evaluation as a management tool. Finally, this phase builds the foundation for both process and impact evaluations.

The purpose of the "**process evaluation**" is to provide evaluators with technical information that can assist it in developing and implementing similar programs. Special attention is given to identifying the lessons learned at the various sites and the guidance they can provide to other jurisdictions regarding organization and development of program activities and to the implementation of program elements which are useful in addressing particular kinds of purposes or problems. Notably, performing process evaluations establishes the foundation for more intensive impact evaluations.

By determining the effectiveness of program activities, future program options can be considered and decisions on program design can be made. Process evaluation permits the initial comparison of objectives and performance measures; its products represent essential interim results that help keep programs on track, or even help to halt activities that are having little effect and redirect limited resources elsewhere.

The purpose of the "**impact/intensive evaluation**" is to provide management information needed by Federal, State or local officials and community leaders involved in policy and programming decisions which clearly confirms that specific programs and/or activities do work, or do not work. The impacts observed in the demonstration projects are distilled to provide assessment of the impact to the program strategies.

De-mystifying Assessment and Evaluation

The term "**program**" refers to an organized set of activities that are managed toward a particular set of goals for which the program can be held separately accountable. Evaluations can be directed at determining overall program effectiveness (e.g., an entire youth anti-violence strategy) or the effectiveness of particular program components, or sub-programs (e.g., the alternatives to corrections program). It is not always feasible to do the former, due to the level of generality involved. An alternative is to focus on program components (the parts) as a means of evaluating the overall program (the whole). Within a program, some components may be well defined and properly implemented, while others are not. Therefore, a pertinent question for program managers is what aspects, or components, of the program should be considered for inclusion in an assessment or evaluation.

What is needed before implementation? First, a "program description" is always useful, and can in itself form the basis for the remaining steps of the study. The identification of goals

and objectives is not as easy a task as it may appear, but a comprehensive program description helps a great deal.

A "**Goal**" is the end toward which effort is directed. Program goals should be clearly stated and realistic. "**Objectives**" are the intermediate effects or results to be achieved by the program in pursuing its ultimate goal. Objectives represent the means by which program managers determine the extent program goals are being accomplished. It is imperative that objectives be both achievable and measurable. Like goals, objectives should be stated in terms of outcomes (expected effects or results).

The most important step to understanding a program's implementation, however, is the knowledge gained by identifying "**program activities**" that are expected to produce results which meet the stated goals and objectives. The program must have a realistic chance of attaining its specified goals and objectives, if it is to be implemented. Therefore, the cause-and-effect relationship between program activities and goals and objectives must be identified and assessed.

To be effective, evaluation measurements should focus on the lowest possible level of the program flow model (on program activities versus goals/objectives). Activities are more specific and well-defined than goals/objectives and thus allow better measures of program effectiveness. Therefore, evaluations should focus on program activities as a strategy for assessing projects or the larger programs of which they are a part. This strategy provides the basis for documenting, modifying and/or eliminating particular program activities in response to identified weaknesses in the program.

Measuring the impact of activities in relation to the goals and objectives which they seek to achieve requires the development of "**performance indicators**." An **indicator** is defined as an explicit measure of effects or results expected. It tells to what extent an activity has been successful in achieving, or contributing to, an objective.

Indicators may be quantitative or qualitative. A quantitative indicator can be expressed as a single measure (number of individuals on parole), or as a degree of change (increase/ decrease in number of domestic violence cases). Qualitative indicators can be used where quantitative measures are not feasible. It is not possible, for example, to assign a direct quantitative measure to the extent to which neighborhoods have been made safer through crime watch programs. However, a qualitative (or indirect) measure can be used through the use of surveys, direct observation, etc.

The above analytical model demonstrates the linear progression of evaluation and performance data from program goals to specific program activities, through performance measurement, to analysis and interpretation of results and necessary corrective actions. Information of this type is essential for determining if programs are on track and working well.

Some Common Lessons Learned

First, programs seem to be making a difference where there is collaboration and cooperation.

Second, rather than expecting to see large statewide reductions in violence or the availability of drugs, we should be looking for impact at the program or neighborhood level.

Third, while evaluation is an important program development tool, it is not an isolated event that can be performed once and then forgotten. Rather, evaluation should be part of a feedback loop that guides program development and operation on an ongoing process.

Finally, while programs and resources will always be diverse, every program can engage in evaluation to some degree. The key is to recognize the efficacy of evaluation. Properly designed and executed evaluations are neither disruptive nor dangerous. They are the program manager's best vehicle for documenting success and getting the necessary feedback to identify and resolve problems. If stakeholders can objectively demonstrate their program's effectiveness, they will be in a better position to compete for limited resources, and also make changes that will help serve both their clients and the public.

A final note.

This discussion is based on a conviction that assessment and evaluation must be considered at the earliest stages of program development. It also pleads the case that there is no reason that useful evaluation methods and approaches cannot be applied to every program and by every program manager. Future discussions, however, must focus directly on the public sector's needs to support and promote long-term research and longitudinal evaluations. If we successfully build the foundations in our programs, efforts to access comprehensive knowledge confirming what works will also be successful.

Appendix A

HOW EVALUATION FITS IN

Program Strategy (Goals)
Program Development (Objectives)
Program Implementation (Activities/Tasks)
Monitoring, Assessment and Evaluation (Feedback and Performance Measures)
Management and Analysis Reports
Review and Adjustment of Strategy, Development & Implementation

COMPLETE EVALUATION SYSTEM

1. Nature and Extent of Current Situation
   (Individual, Group, Specific Location)

2. Program Effectiveness Model
   (Who What How, under what condition)

3. Performance Indicators

4. Program Analysis
   (Compare Expected Versus Actual)

5. Outcome Indicators

6. Program Impact Model
   (When Where Why)

7. Confirmation Criteria

Appendix B-2

GENERAL APPROACHES TO BUILDING AN EVALUATION SYSTEM	
APPLIED, POLICY ORIENTED APPROACH	**THEORY-/RESEARCH-BASED APPROACH**
Goals and Objectives	Prior Research Desired Outcomes Predicting Results of Interventions
Program Logic Design	Hypothesis Testing
Monitoring Progress in Meeting Objectives	Statistical Analysis of Changes in Indicators
Program Analysis of What Is and is NOT Working	Determine If Significant Difference vs. Control
Apply Results to Policy Analysis: If working, does program achieve results compared to other alternatives?	Confirmation under all conditions vs. need for more replication
Promising Criminal Justice Program ready for Replication	Model Criminal Justice Program combined with Longitudinal Research Program

Appendix B-3

INCREMENTAL PHASES OF COMPLETE EVALUATION SYSTEM		
Program Logic Analysis	Establish Linkage between Objectives & Program Activities Determine Desired Outcome and Performance Measures	Management Tool Initial Assessment & Implementation Studies Builds Foundation for Process and Impact Evaluations
Process Evaluation	Determining Effect	Permits Comparison of Objectives and Performance Measures Options/Decisions on Program Design and Success/Failure
Impact/Intensive Evaluation	Determining Impact: -Outcome -Causality Confirmed through Successful Replication	Compares Theory with Outcomes Indicators Options/Decisions for Model Programs, further Replication or Longitudinal Studies

Appendix C-1

PROGRAM EVALUATION: A REALISTIC FRAMEWORK[5]

I. Setting the Foundation

1. Identify problem or need, and current situation.
2. Describe the intervention or program. Identify its goals and objectives (in light of the problem).
3. State the theory underlying the program: How does the program address the problem? Why should the program be effective in achieving the identified goals and objectives?

II. Assessing Program Integrity: Process Evaluation

4. Describe the program as it is implemented. Does the program operate as it is intended? Are the program activities - as they are implemented - likely to achieve the stated goals and objectives?
5. Identify performance indicators that reflect objectives. Assess program integrity, "evaluability."

III. Assessing Program Outcomes: Impact Evaluation

6. Identify intermediate and long-term outcomes. Identify, develop outcome measures.
7. Develop and implement research design; test program impact, effectiveness. Does the program achieve its goals?
8. Inform stakeholders about findings. Identify recommendations; disseminate to wider audience.

---

[5] Appendix C is adapted from an earlier version developed by Douglas Young, Senior Research Associate, Vera Institute of Justice. The example he used was based on the Drug Treatment Alternatives to Prison (DTAP) Program evaluation.

Appendix C-2

I. Setting the Foundation

Task	*Evaluator*	*Examples*
1. Identify problem	^^ ++ can place problem in national context, assess whether others face same problem	>overloaded criminal justice system >alleviate pressure on courts and corrections >enhance intensive supervision >improve treatment services
2. Identify program Goals and objective	^^^ ++ help program designers disaggregate aims, limit and prioritize goals & objectives ++ serve as objective judge concerning the probability of reaching goals -- may overemphasize measurement or force goals to fit preferred research design, risk loss of true program goals, objectives	>establish development & planning group/ identify program participants >identify clients and conduct assessments >achieve, maintain program integrity >reduce rearrest, crime levels >achieve criminal justice system cost savings
3. State the theories that underlie the intervention	^^^ ++ help program designers examine assumptions, confront leaps of faith and faulty reasoning ++ provide proven theories from past literature, judge proposed theory in light of past research -- may force abstract, complex theories and lose sight of more accurate, program driven theories	>criminal behavior is linked to substance abuse >treating substance abuse problem will lead to less crime, lower recidivism >educate/train practitioners >case management approach assures appropriate interventions and increased completion rates >client-designed programs can enhance retention

KEY: ^=pretty useless; ^^=could be of some help; ^^^=can help a lot; ^^^^=a necessity // ++positive role; -- negative role

II. Assessing Program Integrity: Process Evaluation

Task	*Evaluator*	*Examples*
4a. Describe, assess process of program implementation; identify barriers to implementation **4b.** Describe program *as implemented* and compare with original plan & design; test logic model	^^ ++ provides objectivity, credibility ++ can place implementation in national context, compare with prior experience ++ can assess generalizability of program and help describe program for purposes of replication in other locales, systems -- risk divorcing research from operations with "formal" assessment	>frequent meetings between program participants to monitor, resolve implementation issues >solicited program documentation from practitioners >issue annual implementation and operations reports >initiate contract with outside evaluator >evaluator issues various reports
5. Identify and measure performance indicators	^^^ ++ provides objectivity, credibility ++ provides expertise on validity and reliability of quantitative measures ++ lends knowledge about related programs' use of indicators ++ can assess availability of existing indicators, cost of creating and implementing new ones -- may encourage use of overly complex, existing measures instead of indicators that flow from program goals, activities -- overemphasis on quantitative measures may risk loss of more appropriate ones	>identified quantitative performance goals that can be measured: 100+ admissions; 70% retention @ one-year post-admission; <10% abscond and remain at large: >90% of failures receive incarceration; 80% have fewer drug & alcohol charges; <30% recidivism after program completion @ one-year >arrange for program staff to collect data >integrated information management system linking operations, selection, tracking, accomplishments and performance results

: ^=pretty useless; ^^=could be of some help; ^^^=can help a lot; ^^^^=a necessity // ++positive role; -- negative role

III. Assessing Program Outcomes: Impact Evaluation

Task	*Evaluator*	*Examples*
6. Identify intermediate and long-term outcomes and measures	^^^ ++ provides objectivity, credibility ++ expertise on measurement ++ knowledge about related evaluations -- overemphasis on quantitative, overly complex measures	>program retention, completion >post-program criminal recidivism >other post-program outcomes: drug use, employment, housing & family stability, medical and mental health >cost savings due to prison bed displacement , reduced drug use, recidivism
7. Develop and implement research design; test program impact, effectiveness	^^^^ ++ provides expertise on research designs and use of prior, related designs -- may perform impact evaluation prematurely, without foundations -- encourage use on overly complex, inappropriate designs at the expense of assessing true program goals	>developed, implemented experimental or quasi-experimental design for impact evaluation >developed, implemented study on role of specific demonstration vis-a-vis other alternatives >developed, implemented cost-benefit study
8. Inform stakeholders about findings; identify and implement recommendations; disseminate findings	^^^^ ++ can place findings in national context, prior related research ++ identify and reach national audience -- reports may be too technical, inaccessible to stakeholders, practitioners	>regular meetings with program stakeholders, administrators >programs implemented some preliminary recommendations while waiting final outcomes >papers and reports: published, conferences, workshops - disseminate to wide audience

: ^=pretty useless; ^^=could be of some help; ^^^=can help a lot; ^^^^=a necessity // ++positive role; -- negative role